

A classification of invariant distributions and convergence of imprecise Markov chains

Damjan Škulj

February 26, 2013

Abstract

We analyse the structure of imprecise Markov chains and study their convergence by means of accessibility relations. We first identify the sets of states, so-called *minimal permanent classes*, that are the minimal sets capable of containing and preserving the whole probability mass of the chain. These classes generalise the *essential classes* known from the classical theory. We then define a class of extremal imprecise invariant distributions and show that they are uniquely determined by the values of the upper probability on minimal permanent classes. Moreover, we give conditions for unique convergence to these extremal invariant distributions.

Keywords imprecise Markov chain, invariant expectation functional, accessibility relation, convergence of imprecise Markov chains

Math. subj. class. (2000) 60J10, 93B35, 47N30

1 Introduction

The term *imprecise Markov chain* denotes a Markov chain whose parameters are not fully determined, but rather some partial information about them is given in terms of models of *imprecise probabilities*. The systematic modelling of Markov chains with partially determined parameters was initiated by Hartfiel and Seneta (see e.g. [Hartfiel and Seneta, 1994, Hartfiel, 1998] and the references therein) under the name ‘Markov set chains’. Their theory is not formally built on the models of imprecise probabilities, such as Walley’s model ([Walley, 1991]), although they develop similar concepts and methods. More formally the models of imprecise probabilities have been involved in the study of Markov chains by the approaches proposed by Škulj [2009], where general models of interval probabilities (see e.g. [Weichselberger, 2001]) are involved, and de Cooman, Hermans, and Quaeghebeur [2009], who introduce the approach with upper and lower expectation functionals.

One of the main considerations of all mentioned approaches are the conditions for unique convergence to an invariant (imprecise) limit distribution. It has been shown that under certain assumptions imprecise Markov chains behave in a similar way as the precise ones, allowing the possibility to generalise concepts, such as accessibility relations [de Cooman et al., 2009], coefficients of ergodicity [Hartfiel, 1998, Škulj and Hable, 2012], and, for the chains with certain absorption, the invariant distributions conditional on non-absorption [Crossman and Škulj, 2010].

The main topic of the present paper are the imprecise Markov chains that do not converge uniquely, which means that the limit imprecise distribution depends on the initial state or distribution over the set of states. In the classical theory this case is relatively easily translated to the case with unique convergence by partitioning the set of states into disjoint *communication classes*, which are then either *essential* or *transient*. Essential classes have the property that the probability mass concentrated within them and in the transient classes leading to them is preserved within the class. In the long term it is then either distributed among states in a stationary way or moves between the states in *cycles*. Transient classes, though, in the long term transfer all the probability mass to the essential classes. The long term behaviour of a precise Markov chain is then described with the proportions of the probability mass within each essential class, that can then be examined, each separately without influencing the others, as either ergodic or cyclic single communication classes. (For more details see e.g. [Seneta, 2006].)

The situation in the case of imprecise Markov chains is considerably more complex. In the precise case the probability mass can only move from one state into another, which is determined by strictly positive transition probability. In the imprecise case, though, it can happen that instead we have a probability interval that has a strictly positive upper and zero lower bound. Therefore, we have to take into account both scenarios – where a transition is possible and where it is not – within the same model (for better clarification see Example 1). This leads to numerous interesting new possibilities that are not observed in the precise theory. One immediate consequence is that there are different possible ways to define communication relations between states. *Say for instance, when the upper probability of transition from a state x to another state y is positive and the lower one zero, can we say that y is accessible from x or not?* We argue that questions of this kind cannot be answered with a single accessibility relation. Therefore, in addition to the accessibility relation defined by de Cooman et al. [2009], that is based on upper probabilities and which we describe in Subsection 3.1, we define another accessibility relation in Subsection 3.2. This relation is defined between sets of states rather than between single states. Combining both relations will allow us to define minimal classes of states, called *minimal permanent classes*, that generalise essential communication classes from the classical theory in the sense that they are the minimal sets of states capable of possibly preserving the entire probability mass concentrated within them. However, unlike in the case of essential classes in the classical theory, this property does not prevent the possibility that the mass that can possibly be preserved within one class is entirely or partially transferred elsewhere, that is to other minimal permanent classes. Moreover, these classes are neither necessarily disjoint. In effect, a complex network of interdependence between the classes is possible in general. Our first main result in Theorem 2 shows that nevertheless all invariant imprecise distributions that possibly concentrate the entire probability mass in one or more minimal permanent classes are uniquely determined by the set of such classes. Further, our second main result in Theorem 3 shows that when probability mass is at least in the limit concentrated in these classes the imprecise distributions over the set of states converge to the unique invariant distributions. These two theorems present the main results of this paper.

In addition to the main theorems, we believe that the analysis of the pair of

accessibility relations contributes to better understanding of imprecise Markov chains in general, describing the complexity of their structure, and on the other hand, the induced minimal permanent classes seem to present the smallest irreducible units, counterparts to the essential communication classes in the classical Markov chains. Although, considerably more apparently irreducible complexity remains in the imprecise case due to complex interdependencies that may be present between the classes.

The paper has the following structure. In the next section we formally describe the model of imprecise Markov chains. In Section 3 we first describe the accessibility relation based on upper transitions, previously defined by de Cooman et al. [2009], which here is referred as the weak accessibility relation, and state the unique convergence theorem, that was one of the main results in the above reference. The so-called strong accessibility relation is then defined in Subsection 3.2, which allows the definition of the so-called permanent classes in Subsection 3.3. In Section 4 we finally use these results to prove our main results in Theorem 2 and Theorem 3.

2 Imprecise Markov chains

Let \mathcal{X} be a non-empty finite set of *states* and $\{X_n\}_{n \in \mathbb{N}}$ a sequence of random variables taking values in \mathcal{X} . The probability distributions of the variables are assumed to be imprecise. Throughout the paper an *imprecise probability (distribution)* denotes a closed convex set of expectation functionals¹ E . The elements of this set will be denoted by P . Sets of expectation functionals are of course equivalent to sets of corresponding probability distributions. Instead of the term imprecise probability distribution, we will call E an *imprecise expectation functional (IEF)*².

In our case where the set of states is finite every IEF is a compact set. A convenient way to represent IEFs is via their lower and upper expectation functionals as follows. Let E be an IEF and f a real valued map $\mathcal{X} \rightarrow \mathbb{R}$. Then the *lower* and the *upper expectation* with respect to E are

$$\underline{E}(f) = \min_{P \in E} P(f)$$

and

$$\overline{E}(f) = \max_{P \in E} P(f).$$

We will call real valued maps $f: \mathcal{X} \rightarrow \mathbb{R}$ *gambles*. The set of all gambles on \mathcal{X} is denoted by $\mathcal{L}(\mathcal{X})$, or simply by \mathcal{L} . The values of $P(f)$ for a given gamble f and all functionals $P \in E$ therefore form the interval $[\underline{E}(f), \overline{E}(f)]$, and we may therefore identify $E(f)$ with this interval. In particular, when $f = 1_A$, where A is some set of states, the value $P(1_A)$ denotes the probability of A , and $\underline{E}(1_A)$ and $\overline{E}(1_A)$ are the *lower* and the *upper probability* of A .

¹The notation here is very simplified version of the notation used by Walley [1991], which is still enough general and clear for our purpose.

²Imprecise expectation functionals correspond to *lower* and *upper previsions* which constitute one of the key objects studied in the theory of imprecise probabilities. The way how imprecise expectation functionals are used here corresponds to *credal sets* of lower and upper previsions.

The set of IEFs is ordered with respect to set inclusion, and in the case where $E \subseteq F$ we say that E is *more committal*³ than F . Clearly, an IEF E is more committal than F if and only if $\underline{E}(f) \geq \underline{F}(f)$, or equivalently, $\overline{E}(f) \leq \overline{F}(f)$, for every gamble f .

The subset of gambles mapping \mathcal{X} to the interval $[0, 1]$ will be denoted by \mathcal{L}_1 , and it is a compact set. It can be easily verified that the values of imprecise expectation functionals on \mathcal{L}_1 uniquely determine their values on the whole \mathcal{L} . Particularly, $\overline{E}(f) \leq \overline{F}(f)$ for every gamble $f \in \mathcal{L}_1$ implies $E \subseteq F$.

The least committal imprecise expectation functional on $\mathcal{L}(\mathcal{X})$ is the *vacuous* IEF which contains all expectation functionals. The vacuous IEF will usually be denoted by V . For every gamble $f \in \mathcal{L}(\mathcal{X})$ we have

$$\underline{V}(f) = \min_{x \in \mathcal{X}} f(x) \quad \text{and} \quad \overline{V}(f) = \max_{x \in \mathcal{X}} f(x).$$

In the sequel we will list some basic properties of lower and upper expectation functionals. First we have that

$$\underline{E}(f) = -\overline{E}(-f).$$

All the properties of lower expectation functionals are therefore easily derived from the corresponding properties of the upper expectation functionals, which are the following. Let E be an IEF, f, f_1, f_2 arbitrary gambles, λ a non-negative real constant, μ an arbitrary real constant, and 1_A the characteristic function of a set $A \subseteq \mathcal{X}$. Then:

- (i) $\min_{x \in \mathcal{X}} f(x) \leq \underline{E}(f) \leq \max_{x \in \mathcal{X}} f(x)$ (boundedness);
- (ii) $\underline{E}(f_1 + f_2) \leq \underline{E}(f_1) + \underline{E}(f_2)$ (subadditivity);
- (iii) $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ (non-negative homogeneity);
- (iv) $\underline{E}(f + \mu 1_{\mathcal{X}}) = \underline{E}(f) + \mu$ (constant additivity);
- (v) if $f_1 \leq f_2$ then $\underline{E}(f_1) \leq \underline{E}(f_2)$ (monotonicity).

Conversely, every functional satisfying conditions (i)-(iii) above is an upper expectation functional for some IEF.

We will often need the following simple result.

Lemma 1. *Let P be an expectation functional, f a gamble and $a \geq 0$ a constant. Then*

- (i) *if $f \geq 0$ then $P(f) \geq aP(1_{\{f \geq a\}})$;*
- (ii) *$P(f) = \bar{f}$ if and only if $P(1_{\{f = \bar{f}\}}) = 1$,*

where $\bar{f} = \max_{x \in \mathcal{X}} f(x)$. The above inequalities are also valid for \underline{E} and \overline{E} , where E is an IEF.

Proof. To see (i) note that $a1_{\{f \geq a\}} \leq f$ and $aP(1_{\{f \geq a\}}) = P(a1_{\{f \geq a\}}) \leq P(f)$ because of monotonicity of P . (ii) is obvious. \square \square \square

³This expression is a consequence of the fact that a more committal imprecise expectation functional is 'less imprecise', which in Walley's theory commits to accepting more gambles.

We continue with the description of imprecise Markov chains. The imprecise probability distribution corresponding to the random variable X_n is described by an IEF E_n . Given the distribution of X_n , the distribution of X_{n+1} (and consequently of the variables corresponding to later steps) can be found via *transition probabilities*:

$$P(X_{n+1} = \cdot | X_n = x). \quad (1)$$

When the above transition probability is imprecise, it will be denoted with an IEF $T(\cdot|x)$. Given a gamble f the conditional IEF $T(\cdot|x)$ maps it to the set $T(f|x)$. With a fixed gamble the latter expression is an interval mapping on the set of states \mathcal{X} , denoted by Tf . Every $Tf(x) = T(f|x)$ denotes a closed interval $[\underline{T}f(x), \overline{T}f(x)] = [\underline{T}(f|x), \overline{T}(f|x)]$. Hence, Tf is the set of gambles h satisfying $\underline{T}f(x) \leq h(x) \leq \overline{T}f(x)$. The so defined operator T is called an *imprecise transition operator (ITO)*.

An ITO is a closed convex set of (precise) transition operators; however, not every such set is an ITO. (We will denote precise transition operators with small letters, usually t .) It has to be additionally required that a closed convex set of transition operators has *separately specified rows* to be an ITO, which means that if the transition operators t and t' belong to T then for every $x \in \mathcal{X}$ the operator t'' such that

$$t''f(y) = \begin{cases} tf(y) & \text{if } y \neq x \\ t'f(y) & \text{if } y = x \end{cases}$$

also belongs to T . It is easy to verify that separately specified rows of an ITO ensure that there is some $t \in T$ for every $h \in Tf$ so that $h = tf$, and in particular, that $\underline{T}f$ and $\overline{T}f$ belong to Tf .

In order to allow calculations with imprecise transition operators, we must extend the definition of Tf from gambles to intervals of gambles of the form

$$[\underline{f}, \overline{f}] := \{h: \underline{f} \leq h \leq \overline{f}\}.$$

We define

$$T[\underline{f}, \overline{f}] = [\underline{T}f, \overline{T}f],$$

which clearly contains exactly all the gambles of the form tf where $t \in T$ and $f \in [\underline{f}, \overline{f}]$.

Let an imprecise Markov chain be given in terms of an imprecise initial distribution in the form of IEF E_0 , and an ITO T . The imprecise distribution at time n , given in the form of IEF E_n , is calculated with

$$E_n(f) = E_0(T^n f), \quad (2)$$

where $T^n f$ is given recursively with $T^0 f = f$ and $T^{k+1} f = T[T^k f]$ for every $k \geq 0$.

The following holds.

Proposition 1. *Let IEFs E and F be given. Then*

$$(\alpha E + \beta F)T = \alpha ET + \beta FT$$

if $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$.

Proof. Take any gamble f , and calculate

$$\begin{aligned}
(\alpha E + \beta F)(Tf) &= (\alpha E + \beta F)([\underline{T}f, \overline{T}f]) \\
&= [(\alpha \underline{E} + \beta \underline{F})(\underline{T}f), (\alpha \overline{E} + \beta \overline{F})(\overline{T}f)] \\
&= [\alpha \underline{E}(\underline{T}f) + \beta \underline{F}(\underline{T}f), \alpha \overline{E}(\overline{T}f) + \beta \overline{F}(\overline{T}f)] \\
&= \alpha [\underline{E}(\underline{T}f), \overline{E}(\overline{T}f)] + \beta [\underline{F}(\underline{T}f), \overline{F}(\overline{T}f)] \\
&= (\alpha ET + \beta FT)(f).
\end{aligned}$$

□

□

Much of the attention in this paper is put to invariant imprecise expectation functionals, which generalise the concept of invariant distributions. Let T be an ITO. An IEF E is called *T-invariant* if $ET = E$. Clearly, if E is T -invariant, then it is also T^n -invariant, for every $n \in \mathbb{N}$.

It has been shown previously in [Škulj, 2009] that there always exists the least committal T -invariant IEF for every ITO T , and under certain assumptions also the most committal one. We have the following proposition.

Proposition 2. *Let an ITO T be given and an IEF E_0 be such that either*

$$E_0 T \subseteq E_0$$

or

$$E_0 T \supseteq E_0.$$

Then the sequence $\{E_n\}_{n \in \mathbb{N}}$, where $E_{n+1} = E_n T$, is monotone and the limit $E_\infty = \lim_{n \rightarrow \infty} E_n$ exists and is a T -invariant IEF.

It is an immediate consequence of this proposition that taking E_0 to be the vacuous IEF on the set of states \mathcal{X} , the limit IEF is the unique least committal T -invariant IEF, but in general there is no unique most committal T -invariant IEF.

Definition 1. Let E be an IEF. Then we define the *support* of E with

$$\text{supp}(E) = \{x: \overline{E}(1_{\{x\}}) > 0\}.$$

The following proposition is straightforward.

Proposition 3. *For every IEF E and every gamble $f \in \mathcal{L}(\mathcal{X})$ we have that*

$$E(f) = E(f \cdot 1_C),$$

whenever $\text{supp}(E) \subseteq C$.

3 Accessibility relations and convergence

In the theory of Markov chains it is an important question which states are possible to be visited at following times when a state has been observed at given time. In the case of precise Markov chains there is a clear meaning of the expression 'possible', which states that the probability is strictly positive. While

in the case of imprecise Markov chains the situation is more complex. The probabilities are now in the form of intervals, and these intervals can have positive upper bound but zero lower bound. This suggests that there may be different notions of accessibility relations, reflecting different notions of 'possibility'.

We first describe the notion of accessibility introduced by de Cooman et al. [2009], based on upper probability. This accessibility notion is based on proclaiming an event possible if it has strictly positive upper probability. Note that the probability (interval) of the chain being in a state y at time n given that it is in the state x at time 0 is obtained using formula (2) as $T^n 1_{\{y\}}(x)$.

3.1 Weak accessibility relation and unique convergence

The accessibility relation based on the upper probabilities is defined as follows. A state y is *accessible* from x in n steps if $\overline{T}^n 1_{\{y\}}(x) > 0$. We will then write $x \overset{n}{\rightsquigarrow} y$. When $x \overset{n}{\rightsquigarrow} y$ for some $n \geq 0$, we write $x \rightsquigarrow y$. If both $x \rightsquigarrow y$ and $y \rightsquigarrow x$ then we say that the states *communicate* and write $x \rightsquigarrow\!\!\!\rightsquigarrow y$. The relation \rightsquigarrow is a preorder, i.e. is reflexive and transitive, and $\rightsquigarrow\!\!\!\rightsquigarrow$ is an equivalence relation that partitions the set of states \mathcal{X} into *communication classes*. Every two elements of a communication class are then accessible from one another.

Definition 2. A set $C \subseteq \mathcal{X}$ of states is *absorbing* if $x \in C$ and $x \rightsquigarrow y$ implies $y \in C$.

Proposition 4. Let C be absorbing. Then

$$V_C T \subseteq V_C \quad (3)$$

where V_C is the vacuous IEF on C .

Proof. Clearly we have that $\text{supp}(T(\cdot|x)) \subseteq C$ for every $x \in C$, and therefore, by Proposition 3, $\overline{T}f(x) = \overline{T}(f|x) = \overline{T}(f1_C|x) \leq \overline{V}_C(f)$. Hence, $\overline{V}_C(\overline{T}f) \leq \overline{V}_C(f)$, which implies (3). \square

The following corollary follows immediately from Propositions 2 and 4.

Corollary 1. Let C be absorbing. Then there exists the unique least committal T -invariant IEF E such that $\text{supp}(E) \subseteq C$.

Proposition 5. Let E be a T -invariant IEF. Then $\text{supp}(E)$ is absorbing.

Proof. Let $x \in \text{supp}(E)$ and $x \rightsquigarrow y$. By definition then there is some $r > 0$ such that $\overline{T}^r 1_{\{y\}}(x) > 0$. T -invariance of E implies

$$\overline{E}(1_{\{y\}}) = \overline{E}(\overline{T}^r 1_{\{y\}}) \geq \overline{E}(1_{\{x\}}) \overline{T}^r 1_{\{y\}}(x) > 0.$$

Hence $y \in \text{supp}(E)$ as well. \square

Take two communication classes \mathcal{C} and \mathcal{D} . If for some states $x \in \mathcal{C}$ and $y \in \mathcal{D}$, x leads to y , and hence every element in \mathcal{C} leads to every element in \mathcal{D} , then we say that the class \mathcal{C} leads to \mathcal{D} . If a class \mathcal{D} only leads to itself, then we call it a *maximal class*. When there is a unique maximal class it is called the *top class*. A communication class is called *regular* if for every pair of its elements x and y there is some r such that $x \overset{n}{\rightsquigarrow} y$ for every $n \geq r$. If the top

class is regular, then the accessibility relation is said to be *top class regular*. An imprecise Markov chain is called *regularly absorbing* if it is top class regular and if there is some $n \in \mathbb{N}$ such that

$$\underline{T}^n 1_R(y) > 0 \text{ for every } y \in \mathcal{X} \setminus R, \quad (4)$$

where R is the top class. In other words, the lower probability that the chain will eventually move from any class outside the top class to the top class is positive. This practically means that no matter what is the initial distribution of the probability mass, it gets transferred to the top class, making it an absorbing class. So in the long run all probability mass is moved into the top class.

The following convergence theorem is one of the most general results on unique convergence for imprecise Markov chains.

Theorem 1 ([de Cooman, Hermans, and Quaeghebeur, 2009], Theorem 5.1). *Consider a stationary imprecise Markov chain with finite set \mathcal{X} that is regularly absorbing. Then for every initial upper expectation \overline{E}_0 the upper expectation $\overline{E}_n = \overline{E}_0 \overline{T}^n$ for the state at time n converges point-wise to the same upper expectation \overline{E}_∞ :*

$$\lim_{n \rightarrow \infty} \overline{E}_n(f) = \lim_{n \rightarrow \infty} \overline{E}_0(\overline{T}^n f) =: \overline{E}_\infty(f) \text{ for every } f \in \mathcal{L}(\mathcal{X}).$$

Moreover, the limit expectation \overline{E}_∞ is the only \overline{T} -invariant upper expectation on $\mathcal{L}(\mathcal{X})$.

The main objective of this paper is to explore the situations where the assumptions of the above theorem are not satisfied. Let us first give a motivating example with a top class regular imprecise Markov chain that is not regularly absorbing, i.e. does not satisfy (4).

Example 1. We consider an imprecise Markov chain with the set of states \mathcal{X} partitioned in two subsets \mathcal{X}_1 and \mathcal{X}_2 and ITO T that can be represented in the block matrix form:

$$T = \begin{bmatrix} P & \mathbf{0} \\ R & Q \end{bmatrix},$$

where the blocks correspond to \mathcal{X}_1 and \mathcal{X}_2 respectively, so that P is a regular imprecise operator, $\overline{R} > 0$ and $\underline{R} = 0$. Clearly, the corresponding chain is top class regular, where \mathcal{X}_1 is its top class; however, it is not regularly absorbing because of $\underline{R} = 0$.

Let us first show that indeed the chain is not uniquely convergent. In fact we will construct invariant IEFs in three different ways, two of them being clearly different, while the equality between the second and the third one is not obvious and will turn out as a consequence of Theorem 2.

The first T -invariant IEF is the one where the probability mass is concentrated in \mathcal{X}_1 , which is an absorbing set. By Theorem 1 starting with any IEF of the block form $(E_0^1, 0)$ the IEFs $(E_n^1, 0) = (E_0^1, 0)T^n$, converge uniquely to a T -invariant $E_\infty = (E_\infty^1, 0)$. Where E_∞^1 is the unique P -invariant IEF.

Now take V to be the vacuous IEF on \mathcal{X} . Since $\underline{R} = 0$ implies $\overline{Q} = 1$, this implies that $\overline{V}(\overline{T}^n 1_{\mathcal{X}_2}) = 1$, as well as $\overline{V}(\overline{T}^n 1_{\mathcal{X}_1}) = 1$ because of $\overline{P} = 1$. But on the other hand the sequence $\{VT^n\}$ converges to some F_∞ , which must then satisfy both $\overline{F}_\infty(1_{\mathcal{X}_1}) = 1$ and $\overline{F}_\infty(1_{\mathcal{X}_2}) = 1$.

Yet another way to find a T -invariant IEF is the following. Let $F' = (0, V_2)$ where V_2 is the vacuous IEF on \mathcal{X}_2 . Consider an ITO of the form

$$T_1 = \begin{bmatrix} P & \mathbf{0} \\ \mathbf{0} & Q^* \end{bmatrix} \subseteq T.$$

For simplicity we may assume that Q^* is regular, which implies that \mathcal{X}_2 is a communication class. The sequence $F'_n = F'T_1^n$ clearly satisfies the conditions of the Proposition 2 and therefore there exists the limit $F'_\infty = \lim_{n \rightarrow \infty} F'_n = (0, F_\infty^2)$. Assuming regularity of Q^* it is then the unique T_1 -invariant IEF of this form. But then $F'_\infty T \supseteq F'_\infty$, which by Proposition 2 again implies that the limit $\bar{F}_\infty = \lim_{n \rightarrow \infty} F'_\infty T^n$ exists and is T -invariant. Moreover, $\bar{R} > 0$ implies that $\bar{F}_\infty(1_{\mathcal{X}_1}) = \bar{F}_\infty(1_{\mathcal{X}_2}) = 1$. Thus the values of \bar{F}_∞ and \bar{F}_∞ coincide on the characteristic functions of both communication classes in \mathcal{X} . It is less obvious though that this implies that they must be equal, as follows from our first main result in Theorem 2.

Although E_∞ and F_∞ are two extremal invariant IEFs, many more can be constructed simply by taking their convex combinations, as easily follows from Proposition 1. It is still an open question, however, whether every non-extremal invariant IEF can be constructed as a convex combination of the extremal ones.

3.2 Strong accessibility relation

The accessibility relation \rightsquigarrow denotes the possibility that one element is accessible from another. Although, it might be that the lower probability of an x leading to a y is zero and the upper one is positive, which clearly can have essentially different implications than the case where both lower and upper probability are positive. Example 1 describes such a case. The weak accessibility relation \rightsquigarrow is therefore not enough to describe all the relevant properties of the behaviour of an imprecise Markov chain.

For this reason we will define another accessibility relation between sets of states as follows.

Definition 3. Let A and B be arbitrary sets of states. We say that A *strongly leads to B in n steps* if $\bar{T}^n 1_B(x) = 1$ for every $x \in A$. We will then write $A \xrightarrow{n} B$. If there is some $n > 0$ such that $A \xrightarrow{n} B$ then we will say that A *strongly leads to B* and write $A \rightarrow B$.

The strong accessibility relation denotes the case where the chain being in any state belonging to A will move to a set of states B with upper probability 1. It should be noted that this does not mean that the chain will move from A to B with certainty, but only that one of possible scenarios, described with the transition probabilities, is also such a certain move. And here is the crucial difference between the precise and imprecise case. While in the precise case a set cannot strongly lead into two disjoint sets, this is perfectly possible in the imprecise case. This fact implies that the interdependencies between classes of states in the imprecise case can be more complex than in the precise case.

Note also that the case where $A \rightarrow B$ is different from the case where $\{x\} \rightarrow B$ for every $x \in A$, although implies it. Consider the following example.

Example 2. Let $\{x\} \xrightarrow{1} \{y\}$ and $\{y\} \xrightarrow{1} \{x\}$. Then $\{x\} \xrightarrow{2k} \{x\}$ for every $k \in \mathbb{N}$ and $\{y\} \xrightarrow{2k+1} \{x\}$ for every $k \in \mathbb{N}$. However, there is no $n \in \mathbb{N}$ such that $\{x, y\} \xrightarrow{n} \{x\}$, whence $\{x, y\} \not\xrightarrow{} \{x\}$.

To help us analyse the properties of the strong accessibility relation \rightarrow we will define the following two functions. Let E be an IEF. Define

$$\psi_E(A) = \begin{cases} 1 & \overline{E}(1_A) = 1; \\ 0 & \text{otherwise} \end{cases}$$

for every $A \subseteq \mathcal{X}$.

Further let T be an ITO and define

$$\Theta_T(A, B) = \begin{cases} 1 & A \xrightarrow{1} B; \\ 0 & \text{otherwise.} \end{cases}$$

The following proposition follows immediately from the definitions.

Proposition 6. *Let E be an IEF and A, B, C, D, A_i for $i = 1, \dots, n$, sets of states.*

- (i) $\Theta_T(A, B) = 1$ if and only if $\psi_{T(\cdot|x)}(B) = 1$ for every $x \in A$.
- (ii) If $A \subseteq B$ then $\psi_E(A) \leq \psi_E(B)$.
- (iii) If $C \subseteq A$ and $B \subseteq D$ then $\Theta_T(A, B) \leq \Theta_T(C, D)$.
- (iv) $\min_{i=1, \dots, n} \Theta_T(A_i, B) = \Theta_T\left(\bigcup_{i=1}^n A_i, B\right)$.

We define the following operations:

$$(\psi_E * \Theta_T)(B) = \max_{A \subseteq \mathcal{X}} \psi_E(A) \Theta_T(A, B)$$

and

$$(\Theta_T * \Theta_{T'})(A, B) = \max_{C \subseteq \mathcal{X}} \Theta_T(A, C) \Theta_{T'}(C, B).$$

The latter expression, when sequentially applied to the same function, allows defining its n -th power with

$$\Theta_T^n = \underbrace{\Theta_T * \dots * \Theta_T}_{n\text{-times}}.$$

Proposition 7. *Let E be an IEF, T and S ITOs and n a positive integer. Then*

- (i) $\psi_{ET} = \psi_E * \Theta_T$;
- (ii) $\Theta_{TS} = \Theta_T * \Theta_S$;
- (iii) $\Theta_{T^n} = \Theta_T^n$;
- (iv) $\psi_{ET^n} = \psi_E * \Theta_T^n$;

(v) $\psi_{[\lim_{n \rightarrow \infty} ET^n]} \geq \lim_{n \rightarrow \infty} \psi_E * \Theta_T^n$ if the limits exist. Particularly, if $ET \subseteq E$, then we have the equality $\psi_{[\lim_{n \rightarrow \infty} ET^n]} = \lim_{n \rightarrow \infty} \psi_E * \Theta_T^n$.

Proof. (i): Let us first show that $\psi_E * \Theta_T \leq \psi_{ET}$. Let $(\psi_E * \Theta_T)(B) = 1$ for some B . By definition, this implies existence of some A such that $\psi_E(A) = 1$ and $\Theta_T(A, B) = 1$. In other words $\overline{E}(1_A) = 1$ and $\overline{T}1_B(x) = 1$ for every $x \in A$. Hence, by Lemma 1 (i), $\overline{E}\overline{T}(1_B) = \overline{E}(\overline{T}1_B) \geq 1 \cdot \overline{E}(1_A) = 1$, which means, by definition, that $\psi_{ET}(B) = 1$.

Now suppose that $\psi_{ET}(B) = 1$ for some B . By definition then $\overline{E}\overline{T}(1_B) = \overline{E}(\overline{T}1_B) = 1$, which by Lemma 1 (ii) is only possible if $\overline{E}(1_{\{\overline{T}1_B=1\}}) = 1$. Then we denote $A = \{x: \overline{T}1_B(x) = 1\}$, which then satisfies $\psi_E(A) = 1$ and $\Theta_T(A, B) = 1$, whence $(\psi_E * \Theta_T)(B) = 1$, and this proves that $\psi_E * \Theta_T \geq \psi_{ET}$.

(ii): By Proposition 6 (i), $\Theta_{TS}(A, B) = 1$ is equivalent to $\psi_{TS(\cdot|x)}(B) = \psi_{T(\cdot|x)S}(B) = 1$ for every $x \in A$. By (i), this is equivalent to $(\psi_{T(\cdot|x)} * \Theta_S)(B) = 1$ for every $x \in A$. Therefore, for every $x \in A$ we have a set C_x such that $\psi_{T(\cdot|x)}(C_x) = 1$ and $\Theta_S(C_x, B) = 1$. Now let $C = \bigcup_{x \in A} C_x$ and, by Proposition 6 (ii) and (iv), we still have that $\psi_{T(\cdot|x)}(C) = 1$ for every $x \in A$ and $\Theta_S(C, B) = 1$. Hence, $\Theta_T(A, C) = 1$ and $\Theta_S(C, B) = 1$, which implies $(\Theta_T * \Theta_S)(A, B) = 1$.

Conversely, if we assume that $(\Theta_T * \Theta_S)(A, B) = 1$, this implies the existence of some C such that $\Theta_T(A, C) = 1$ and $\Theta_S(C, B) = 1$. By Proposition 6 (i), the first equality is equivalent to $\psi_{T(\cdot|x)}(C) = 1$ for every $x \in A$. Therefore $\psi_{T(\cdot|x)S}(B) = 1$ for every $x \in A$; which, as follows from above, is equivalent to $\Theta_{TS}(A, B) = 1$.

(iii) is an immediate consequence of (ii), and (iv) follows from (i) and (iii).

(v) Let $\lim_{n \rightarrow \infty} (\psi_E * \Theta_T^n)(A) = 1$. Then there must exist some n_0 such that $(\psi_E * \Theta_T^n)(A) = 1$ for every $n \geq n_0$ and this is equivalent to $\psi_{ET^n}(A) = 1$. Thus $\overline{E}\overline{T}^n(A) = 1$ for every $n \geq n_0$, which implies that $\lim_{n \rightarrow \infty} \overline{E}\overline{T}^n(A) = 1$ or $\psi_{[\lim_{n \rightarrow \infty} ET^n]}(A) = 1$.

If $ET \subseteq E$ then $ET^n \subseteq ET^{n-1}$ holds for every n . Now suppose that $\lim_{n \rightarrow \infty} \overline{E}\overline{T}^n(A) = 1$. Then the same must hold for every n : $\overline{E}\overline{T}^n(A) = 1$. Hence $\psi_E * \Theta_T^n(A) = 1$ for every n and so is the limit $\lim_{n \rightarrow \infty} \psi_E * \Theta_T^n(A) = 1$. \square \square

Let us give an example where we have strict inequality in Proposition 7 (v).

Example 3. Take the transition matrix

$$T = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix},$$

for a Markov chain with the set of states $\mathcal{X} = \{x_1, x_2\}$. Clearly, starting with the initial probability distribution, say $q_0 = (0.5, 0.5)$, we have $\psi_E(A) = 0$ for all strict subsets $A \subsetneq \mathcal{X}$. The same holds for every distribution $q_n = q_0 T^n$, but the limit distribution is $(1, 0)$. Hence, we have $\psi_{q_n}(A) = 0$ for every strict subset of \mathcal{X} , and the same for the limit $\lim_{n \rightarrow \infty} \psi_{q_n}(A)$. However, $\psi_{[\lim_{n \rightarrow \infty} q_n]}(\{x_1\}) = 1$.

Corollary 2. Let A and B be sets of states such that $A \xrightarrow{n} B$. Then a sequence of sets

$$A = A_0, A_1, \dots, A_n = B \tag{5}$$

exists such that

$$A_i \xrightarrow{1} A_{i+1}, \text{ for every } i = 0, \dots, n-1. \quad (6)$$

Proof. The relation $A \xrightarrow{n} B$ is equivalent to $\Theta_{T^n}(A, B) = 1$, and by Proposition 7(iii), this is again equivalent to $\Theta_T^n(A, B) = 1$. The existence of the sequence is now an immediate consequence of the definitions. $\square \quad \square$

3.3 Permanent classes

Definition 4. A set A of states is called a *permanent class* if for every $N \in \mathbb{N}$ an $n \geq N$ and a set B exist such that $B \xrightarrow{n} A$. Any set of states that is not permanent will be called an *impermanent class*.

In the case of precise Markov chains only absorbing classes are permanent, while the imprecise case allows much richer structure.

Example 4. Consider the transition operator from Example 1. Clearly, the disjoint subsets of states \mathcal{X}_1 and \mathcal{X}_2 are both permanent classes, despite the fact that \mathcal{X}_2 'possibly' leads to \mathcal{X}_1 .

The following proposition is immediate.

Proposition 8. *Every subset of states A such that $A \rightarrow A$ is a permanent class.*

Definition 5. A permanent class B is *minimal* if $B \rightarrow B$ and contains no proper subset $B' \subset B$ such that $B' \rightarrow B'$.

We now examine some properties of permanent classes. First we give some equivalent definitions.

Proposition 9. *Let a set of states A be given. The following propositions are equivalent:*

- (i) A is a permanent class.
- (ii) For every $N \in \mathbb{N}$ an $n \geq N$ exists and a sequence of sets $A_0, A_1, \dots, A_n = A$ such that $A_i \xrightarrow{1} A_{i+1}$ for every $i = 0, \dots, n-1$.
- (iii) For every $n \in \mathbb{N}$ a sequence of sets with the above properties exists.
- (iv) A permanent class B exists such that $B \rightarrow A$.
- (v) A minimal permanent class B exists such that $B \rightarrow A$.

Proof. The equivalence between (i), (ii) and (iii) clearly follows by Corollary 2. Further, (iv) and (v) clearly imply (iii).

Now let A be a permanent class. Then for every n we have a sequence $A_{n,0}, \dots, A_{n,n-1}$ such that $A_{n,i} \xrightarrow{1} A_{n,i+1}$ for every n and $i < n$ and $A_{n,n-1} \xrightarrow{1} A$. Because of finiteness, there exists a set B such that $B = A_{n,n-1}$ for an infinite number of indices n , and this implies that B satisfies (ii) and is therefore a permanent class. Thus, (i) implies (iv) as well.

To see that (iv) implies (v), let A be a permanent class and $B_1 \rightarrow B_2 \rightarrow \dots \rightarrow B_r = A$ a maximal chain such that $B_{k+1} \not\rightarrow B_k$ for every $k = 1, \dots, r-1$.

By (iv) there exists a permanent class C such that $C \rightarrow B_1$. But then, by maximality of the chain, we must have that also $B_1 \rightarrow C$, and consequently, by transitivity, $B_1 \rightarrow B_1$. If there is a strict subset $B' \subset B_1$ that is a permanent class then we take it to be a minimal one. We have then $B' \rightarrow B_1$, by construction. But $B_1 \rightarrow B'$, by maximality of the chain. Therefore, by transitivity, we again have that $B' \rightarrow B'$ which is now clearly a minimal permanent class, and such that $B' \rightarrow A$. \square \square

Proposition 10. *Let B be a minimal permanent class. Then it is a communication class. Moreover, there is some $r \in \mathbb{N}$ such that*

$$\overline{T}^r 1_{\{y\}}(x) > 0 \quad \text{for every } x, y \in \mathcal{X} \quad (7)$$

and

$$\overline{T}^r 1_B(x) = 1 \quad \text{for every } x \in B. \quad (8)$$

This means that for some $r \in \mathbb{N}$ the chain with the ITO T^r restricted to B is regular.

Proof. Let us first show that B is a communication class. We have that $B \rightarrow B$, say $B \xrightarrow{n} B$ for some $n \in \mathbb{N}$. Take some $x \in B$ and denote $\tilde{B}_x = \{y \in B : x \not\rightsquigarrow y\}$ and further let $B_x = \{x' \in B : x' \not\rightsquigarrow \tilde{B}_x\}$. We show that $B_x \xrightarrow{n} B_x$. Suppose contrary that there is some $x' \in B_x$ such that $\overline{T}^n 1_{B_x}(x') < 1$. By subadditivity of $\overline{T}^n(\cdot|x')$ we have that $\overline{T}^n 1_{B \setminus B_x}(x') \geq \overline{T}^n 1_B(x') - \overline{T}^n 1_{B_x}(x') > 0$. Hence, there must be some $y \in B \setminus B_x$ such that $\overline{T}^n 1_{\{y\}}(x') > 0$, or equivalently, $x' \xrightarrow{n} y$. But since $y \notin B_x$, $y \rightsquigarrow \tilde{B}_x$, which together implies that $x' \rightsquigarrow \tilde{B}_x$, which is in contradiction with $x' \in B_x$.

It is therefore confirmed that $B_x \rightarrow B_x$, which contradicts minimality of B , unless $B_x = \emptyset$. Together this proves that B is a communication class.

Let us now show that some $r \in \mathbb{N}$ exists so that equations (7) and (8) are satisfied. Let $x \in B$ and denote with r_x the greatest common divisor of all r such that $x \xrightarrow{r} x$. Now denote $C_x = \{y \in B : \exists k, x \xrightarrow{kr_x} y\}$.

Since $\overline{T}^{kr_x} 1_{B \setminus C_x}(y) = 0$ for every $y \in C_x$, which implies that $\overline{T}^{kr_x} 1_{B_x}(y) = 1$ for every $y \in C_x$. Thus, $C_x \xrightarrow{kr_x} C_x$ and therefore, $C_x \rightarrow C_x$. Since B is a minimal permanent class and C_x a non-empty subset, it must therefore hold that $C_x = B$ for every $x \in B$.

We now show that there is some $r \in \mathbb{N}$ such that $\overline{T}^r 1_{\{y\}}(x) > 0$ for every $x, y \in B$. It follows from the definition of r_x and the elementary properties of accessibility relations that there is some $L \in \mathbb{N}$ such that $x \xrightarrow{lr_x} x$ for every $l \geq L$, and for every $y \in B$ there exists some l_y such that $x \xrightarrow{l_y r_x} y$. Denote $M_x = \max_{y \in B} (L + l_y)$. Then $x \xrightarrow{mr_x} y$ for every $y \in B$ and every $m \geq M_x$. Now denote $R = \prod_{x \in B} r_x$ and $M = \max_{x \in B} M_x$. Clearly then $x \xrightarrow{mR} y$ for every x and $y \in B$, whenever $m \geq M$.

Because B is a minimal permanent class, we have that $B \xrightarrow{s} B$. If $s = 1$ then (8) is satisfied for every $r \in \mathbb{N}$ and therefore also for $r = mR$. While in the case where $s > 1$ we can replace T with T^s and reason as above to show that $(T^s)^{mR} = T^{smR}$ satisfies both properties required, since, as can easily be noted, B is still a minimal permanent class when T is replaced with T^s . \square \square

Note that the above proposition does not claim that a minimal permanent classes would be whole communication classes, but merely that every minimal permanent class lies in a single communication class.

Definition 6. Let B be a minimal permanent class, such that $B \xrightarrow{r} B$. Then we define

$$T_B^r := \{t \in T^r : t1_B(x) = 1 \text{ for every } x \in B\}. \quad (9)$$

Proposition 11. Let B be a minimal permanent class, such that $B \xrightarrow{r} B$. Then B is regular and absorbing with respect to the weak accessibility relation induced by $(T_B^r)^k$ for some $k \in \mathbb{N}$.

Proof. Let \rightsquigarrow and \rightarrow now denote the weak and the strong accessibility relation induced by T_B^r . By construction we have that $B \rightarrow B$ and since T_B^r is a subset of T^r this relation cannot hold for any proper subset of B . Therefore B is a minimal permanent class with respect to T_B^r as well.

By Proposition 10 it is then a communication class and, moreover, for some $k \in \mathbb{N}$ the power $(T_B^r)^k$ restricted to B is regular. By construction, B is also absorbing with respect to T_B^r and therefore, with respect to $(T_B^r)^k$ as well. \square

4 Unique convergence for extremal IEFs

In this section we use the results prepared in previous sections to state our main results on convergence of imprecise Markov chains. We will explore the convergence of the class of so called extremal IEFs, which are those which induce the upper probabilities of minimal permanent classes being either 0 or 1. In the next subsection we show that when the sequence $\{ET^n\}$ consists of extremal IEFs this sequence converges if the sequence $E(T^n 1_B)$ converges (either to 0 or 1) for every minimal permanent class B , and the limit is uniquely determined by the limits $\lim_{n \rightarrow \infty} E(T^n 1_B)$ for all minimal permanent classes B .

Moreover, in Subsection 4.2 we show that it is sufficient for the unique convergence that the limits $\lim_{n \rightarrow \infty} E(T^n 1_B)$ are zero-one valued, rather than all terms of the sequences.

4.1 Uniqueness of invariant imprecise expectation functionals

The importance of permanent classes is illustrated with the following proposition.

Proposition 12. Let M be the unique least committal T -invariant IEF. Then $\psi_M(A) = 1$ if and only if A is a permanent class.

Proof. The unique least committal IEF M is equal to $\lim_{n \rightarrow \infty} VT^n$, where V is the vacuous IEF on \mathcal{X} . Since $VT \subseteq V$, Proposition 7(v) gives that

$$\psi_M = \psi_{[\lim_{n \rightarrow \infty} VT^n]} = \lim_{n \rightarrow \infty} \psi_V * \Theta_T^n.$$

If $\psi_M(A) = 1$, this implies that for every $n \in \mathbb{N}$ we have that $\psi_V * \Theta_T^n(A) = 1$, or equivalently, there exists some B such that $\psi_V(B) = 1$ and $B \xrightarrow{n} A$. Hence, A is a permanent class.

Moreover, if A is a permanent class, then for every large enough $n \in \mathbb{N}$ there exists some B such that $B \xrightarrow{n} A$, and since $\psi_V(B) = 1$ for every $B \subseteq \mathcal{X}$, then $\psi_V * \Theta_T^n(A) = 1$. As this holds for every large enough integer n , so it does for the limit, and implies that $\psi_M(A) = 1$. \square \square

Definition 7. Let E be an IEF and f a gamble. Then we define *essential maximum* of f given E with

$$\text{ess max}_E f = \max\{a : \underline{E}(1_{\{f \geq a\}}) > 0\},$$

and

$$m_E = \min\{\underline{E}(1_A) : A \subseteq \mathcal{X}, \underline{E}(1_A) > 0\}.$$

Proposition 13. Let E and F be IEFs and f a gamble. The following propositions hold:

- (i) $\underline{E}(f) \leq \text{ess max}_E f$.
- (ii) If $f \geq 0$ then $\underline{E}(f) \geq m_E \cdot \text{ess max}_E f$.
- (iii) If $\psi_E = \psi_F$, then $\text{ess max}_E f = \text{ess max}_F f$.
- (iv) Let $\{E_n\}$ be a sequence of IEFs such that $E_{n+1} \subseteq E_n$. Then $\lim_{n \rightarrow \infty} m_{E_n}$ exists and there is an $N \in \mathbb{N}$ such that $\{m_{E_n}\}_{n \geq N}$ is a non-decreasing sequence.

Proof. Let in this proof \tilde{f} denote $\text{ess max}_E f$.

To see (i), take some $P \in E$ such that $P(1_{\{f > \tilde{f}\}}) = 0$. This means that $P(1_{\{x\}}) = 0$ for every $x \in \mathcal{X}$ such that $f(x) > \tilde{f}$. Then

$$\underline{E}(f) \leq P(f) = \sum_{x \in \mathcal{X}} P(x)f(x) = \sum_{x : f(x) \leq \tilde{f}} P(x)f(x) \leq \tilde{f}.$$

By Lemma 1 (i) and (i) of this proposition, we obtain

$$\underline{E}(f) \geq \tilde{f} \underline{E}(1_{\{f \geq \tilde{f}\}}) \geq \tilde{f} m_E,$$

which proves (ii).

To see (iii) notice that $\psi_E = \psi_F$ if and only if $\underline{E}(1_{\{f > a\}}) > 0$ whenever $\underline{F}(1_{\{f > a\}}) > 0$.

(iv): To see this note that the set $\{\underline{E}_n(1_A) : n \in \mathbb{N}\}$ is a non-decreasing sequence for every $A \subseteq \mathcal{X}$, either constantly equal to 0, or contains a minimal non-zero element. Because of the finite number of subsets of \mathcal{X} we have at most a finite number of positive minima, whose minimum exists and is positive as well.

Clearly, there is also an $N \in \mathbb{N}$ such that for every $A \subseteq \mathcal{X}$ either $\underline{E}_n(1_A) > 0$ for every $n \geq N$ or $\underline{E}_n(1_A) = 0$ for every $n \geq N$. Therefore, $\{m_{E_n}\}_{n \geq N}$ is a non-decreasing sequence. \square \square

Corollary 3. Let E and F be IEFs such that $\psi_E = \psi_F$. Then

$$\underline{E}(f) \geq m_E \cdot \underline{F}(f) \tag{10}$$

for every gamble $f \geq 0$.

Proof. By Proposition 13(ii) we have that $\underline{E}(f) \geq m_E \cdot \text{ess max}_E f$, and by (i) of the same proposition we have that $\underline{F}(f) \leq \text{ess max}_F f = \text{ess max}_E f$. This together implies (10). \square \square

Corollary 4. *Let E and F be IEFs such that $\psi_E = \psi_F$, $\text{supp}(F) \subseteq \text{supp}(E)$, and V_E the vacuous IEF on $\text{supp}(E)$. Then*

$$E \subseteq m_E F + (1 - m_E) V_E.$$

Proof. We need to prove that

$$\underline{E}(f) \geq m_E \underline{F}(f) + (1 - m_E) \underline{V}_E(f) \quad (11)$$

holds for every gamble $f \in \mathcal{L}(\mathcal{X})$. Let f be a gamble and denote $h = f \cdot 1_{\text{supp}(E)}$. Then, by Proposition 3, we have that $H(f) = H(h)$ where H stands for E, F and V_E respectively.

Denote $\underline{h} = \min_{x \in \text{supp}(E)} f(x) = \underline{V}_E(f)$. Then we have that $h - \underline{h} \geq 0$. By Corollary 3 we have that $\underline{E}(h - \underline{h}) \geq m_E \underline{F}(h - \underline{h})$. Hence, by constant additivity of \underline{E} and \underline{F} , $\underline{E}(h) - \underline{h} \geq m_E \underline{F}(h) - m_E \underline{h}$ holds. By replacing \underline{h} with $\underline{V}_E(f)$, $\underline{E}(h)$ with $\underline{E}(f)$ and $\underline{F}(h)$ with $\underline{F}(f)$ we obtain

$$\underline{E}(f) - \underline{V}_E(f) \geq m_E \underline{F}(f) - m_E \underline{V}_E(f),$$

whence inequality (11) follows. \square \square

Definition 8. Let E be an IEF such that $\overline{E}(1_B) \in \{0, 1\}$ for every minimal permanent class B . Then we say that E is an *extremal* IEF.

In the precise case the concept of an extremal expectation functional is rather trivial. Note that a minimal permanent class in the precise case can only be absorbing. Therefore having the entire probability mass once concentrated in such a class it will remain there forever, and additionally, it can only be concentrated in one such class at the time. Consequently, if the class is aperiodic, the probability distributions will then converge to a unique distribution invariant for the chain restricted to this, regular and absorbing class.

Definition 9. Let E be an IEF. Then we define

$$S_E = \{y : \exists x \in \text{supp}(E), x \rightsquigarrow y\}$$

and V_E will from now on denote the vacuous IEF on S_E .

Clearly, the following holds.

Proposition 14. *For every IEF E , the set S_E is absorbing. Moreover, the sequence*

$$M_n = V_E T^n$$

is monotone: $M_{n+1} \subseteq M_n$, and the limit

$$M = \lim_{n \rightarrow \infty} M_n$$

exists and is T -invariant.

Proof. It is an immediate consequence of the definition that S_E is absorbing. The convergence then follows from Propositions 2 and 4. \square \square

Lemma 2. *Let A be a permanent class such that $A \xrightarrow{r} A$ and S an absorbing set. Then $A \cap S \xrightarrow{r} A \cap S$.*

Proof. Clearly, S being absorbing implies that $\overline{T}^r 1_C(x) = 0$ for $C \cap S = \emptyset$ and $x \in S$. Now take some $x \in A \cap S$. By subadditivity of $\overline{T}(\cdot|x)$ we have that $1 = \overline{T}^r 1_A(x) \leq \overline{T}^r 1_{A \cap S}(x) + \overline{T}^r 1_{A \setminus S}(x)$; whence, by $\overline{T}^r 1_{A \setminus S}(x) = 0$, $\overline{T}^r 1_{A \cap S}(x) = 1$ follows. \square \square

Corollary 5. *Let B be a minimal permanent class and S an absorbing set. Then $B \cap S \in \{B, \emptyset\}$.*

Proposition 15. *Let $A \xrightarrow{n} B$ and let E be an IEF. Then $\overline{E}(\overline{T}^n 1_B) \geq \overline{E}(1_A)$.*

Proof. We have that $\overline{E}(\overline{T}^n 1_B) \geq \overline{E}(1_A) \cdot \min_{x \in A} \overline{T}^n 1_B(x) = \overline{E}(1_A)$. \square \square

Corollary 6. *Let $A \rightarrow B$ and let E be a T -invariant IEF. Then $\overline{E}(1_B) \geq \overline{E}(1_A)$.*

Lemma 3. *Let E be an IEF such that ET^n is extremal for every $n \geq 0$ and the limit $\lim_{n \rightarrow \infty} \psi_{ET^n}(B)$ exists for every minimal permanent class B . Then the limit*

$$\psi(A) := \lim_{n \rightarrow \infty} \psi_{ET^n}(A)$$

exists for every $A \subseteq \mathcal{X}$ and ψ is uniquely determined by its restriction to minimal permanent classes.

Proof. Denote $E_n = ET^n$. Let us first show that for every minimal permanent class B $\lim_{n \rightarrow \infty} \psi_{E_n}(B) = 1$ if and only if $B \subseteq S_E$. To see this, take some minimal permanent class $B \subseteq S_E$ and note that $\overline{E}_k(1_B) > 0$ must hold by definition of S_E for some $k \geq 0$. Extremality of E_k implies that $\overline{E}_k(1_B) = 1$, and the fact that $B \rightarrow B$, say $B \xrightarrow{r} B$, by Proposition 15, implies that $\overline{E}_{k+mr}(1_B) = 1$, for every $m \in \mathbb{N}$, whence, by the assumed convergence of the sequence $\{\psi_{E_n}(B)\}$, we must have that, for some $N \in \mathbb{N}$, $\psi_{E_n}(B) = 1$ for every $n \geq N$.

On the other hand, if $B \not\subseteq S_E$, then by Corollary 5, $B \cap S_E = \emptyset$, and therefore clearly, $\overline{E}_n(1_B) = 0$ for every $n \geq 0$.

Now let $A \subseteq \mathcal{X}$. Then either $\psi_{E_n}(A) = 0$ for all sufficiently large n or for every $N \in \mathbb{N}$ there exists an $n \geq N$ such that $\psi_{E_n}(A) = 1$. The former case implies convergence, so let us show that the latter case does as well. Suppose that $\psi_{E_n}(A) = 1$. By Proposition 7 (iv) we have that $\psi_{E_n}(A) = (\psi_E * \Theta_T^n)(A) = 1$, which implies the existence of a sequence $A_0 \rightarrow A_1 \rightarrow \dots \rightarrow A_n = A$, such that $\psi_E(A_0) = 1$. If n is sufficiently large, then at least two members of the sequence are equal, say $A_k = A_l = A'$. Hence, we have $A' \rightarrow A' \rightarrow A$. By Lemma 2 we have that $B' = A' \cap S_E$ also satisfies $B' \rightarrow B'$, and moreover $B' \rightarrow A' \rightarrow A$. However, B' either is a minimal permanent class or contains one. Let B denote such a minimal permanent class, which by construction lies within S_E , and therefore, by the assumptions, an $N \in \mathbb{N}$ exists such that $\psi_{E_n}(B) = 1$ for every $n \geq N$. We have that $B \rightarrow A$, say $B \xrightarrow{r} A$. Moreover, by Proposition 15 we have that $\psi_{E_{r+n}}(A) \geq \psi_{E_n}(B) = 1$ for every $n \geq N$. Hence, $\psi_{E_k}(A) = 1$ for every $k \geq r + N$. \square \square

Theorem 2. *Let E be an IEF such that $E_n = ET^n$ is extremal for every $n \geq 0$ and the limit $\lim_{n \rightarrow \infty} \psi_{E_n}(B)$ exists for every minimal permanent class B . Then the limit*

$$E_\infty := \lim_{n \rightarrow \infty} E_n$$

exists. Moreover, $\lim_{n \rightarrow \infty} E_n = \lim_{n \rightarrow \infty} F_n$ if and only if $S_E = S_F$.

Proof. By Lemma 3 and the assumptions of the theorem, the limit $\psi = \lim_{n \rightarrow \infty} \psi_{E_n}$ exists, and since ψ_{E_n} are discrete valued, there is some $N_1 \in \mathbb{N}$ such that $\psi_{E_n} = \psi$ for every $n \geq N_1$.

It has been shown in the proof of Lemma 3 that $\psi(B) = 1$ for exactly those minimal permanent classes B that are contained in S_E .

Let V_E be the vacuous IEF on S_E and denote $M_n = V_E T^n$ and $M = \lim_{n \rightarrow \infty} M_n$, which exists by Proposition 14. Similar arguments as above show that for every minimal permanent class B we have that $\psi_M(B) = 1$ if and only if $B \subseteq S_E$, whence $\psi_M = \psi$. Moreover, by Proposition 7 (v), $\psi_M = \lim_{n \rightarrow \infty} \psi_{M_n}$, and therefore, there is some $N_2 \in \mathbb{N}$ such that $\psi_{M_n} = \psi_M = \psi$ for every $n \geq N_2$.

Because of $E \subseteq V_E$, we also have that $\text{supp}(E_n) \subseteq \text{supp}(M_n)$ for every $n \in \mathbb{N}$. Now let $N = \max\{N_1, N_2\}$ and denote $m = m_{M_N}$ (c.f. Definition 7). By Corollary 4 we then have that

$$M_N \subseteq mE_N + (1 - m)V_E.$$

Now take any gamble $f \in \mathcal{L}$ and any $n \geq 0$. By Proposition 1 we obtain

$$\overline{M}_N(\overline{T}^n f) \leq m\overline{E}_N(\overline{T}^n f) + (1 - m)\overline{V}_E(\overline{T}^n f).$$

Letting $n \rightarrow \infty$ in the above equation we obtain

$$\overline{M}(f) \leq m \liminf_{n \rightarrow \infty} \overline{E}_n(f) + (1 - m)\overline{M}(f),$$

implying that $\overline{M}(f) \leq \liminf_{n \rightarrow \infty} \overline{E}_n(f)$. But since, clearly, $\overline{M}_n(f) \geq \overline{E}_n(f)$ for every $n \in \mathbb{N}$, this is only possible if the limit

$$\overline{E}_\infty(f) = \lim_{n \rightarrow \infty} \overline{E}_n(f)$$

exists and is equal to $\overline{M}(f)$. Since this holds for every gamble f , we have that $E_\infty = M$. Now since M is uniquely determined by S_E , the limit is the same for all IEFs E that have the same S_E . \square \square

4.2 Unique convergence to extremal invariant IEFs

In the previous subsection we have shown that extremal IEFs converge to a unique invariant IEF that is uniquely determined with the behaviour of the corresponding ψ_E restricted to minimal permanent classes. But often the probability mass is not initially entirely concentrated in a minimal permanent class, but rather it accumulates there as time goes to infinity. It may then happen that in the limit the upper probabilities of all minimal permanent classes are either 0 or 1, although in any finite time they can be somewhere between. In this subsection we show that a convergence to an extremal IEF is unique whenever the upper expectations on minimal permanent classes converge to either 0 or 1. Moreover the limit IEFs are still uniquely determined with the behaviour on the minimal permanent class structure.

Proposition 16. *Let A be an absorbing set. Then $1_A \cdot T^n f = 1_A \cdot T^n[1_A \cdot f]$ for every $n \in \mathbb{N}$.*

Proof. The fact that A is absorbing implies that $T1_{A^c}(x) = 0$ for every $x \in A$, and consequently, $Tf(x) = T[1_A f](x)$ for every $x \in A$. Hence

$$1_A T f = 1_A T[1_A f]. \quad (12)$$

But if A is absorbing for T , then it is also absorbing for every T^n , where $n \in \mathbb{N}$, and therefore the equation (12) holds when T is replaced with T^n . $\square \quad \square$

Proposition 17. *Let A be an absorbing set such that the weak accessibility relation induced by T and restricted to A is regular, and E an IEF such that*

$$\text{supp}(E) \subseteq A. \quad (13)$$

Then

$$M_A := \lim_{n \rightarrow \infty} E T^n$$

exists and is the same for every E satisfying (13).

Proof. Let us first define the following transition operator:

$$T_A f(x) = \begin{cases} T f(x) & x \in A; \\ V_A f & x \notin A, \end{cases}$$

where V_A is the vacuous IEF on A . Clearly, the chain with the ITO T_A is regularly absorbing with the regular top class A . Therefore, by Theorem 1, the sequence $\{E T_A^n\}$ converges uniquely to some M_A . Furthermore, Proposition 16 and the definition of T_A imply that $1_A T_A^n f = 1_A T_A^n[1_A f] = 1_A T^n[1_A f] = 1_A T^n f$ for every gamble f . Using Proposition 3 and the above equalities we obtain

$$E(T^n f) = E(1_A T^n f) = E(1_A T_A^n f) = E(T_A^n f),$$

and since the right hand sides converge, so do the left ones, for every f . $\square \quad \square$

Corollary 7. *Let A be a set of states and $T' \subseteq T$ an ITO such that A is absorbing and regular with respect to the weak accessibility relation induced by T' . Then there exists a T -invariant IEF M_A such that for every IEF E with $\overline{E}(1_A) = 1$ we have that*

$$\liminf_{n \rightarrow \infty} \overline{E}(T'^n f) \geq \overline{M}_A(f) \quad (14)$$

for every gamble f .

Proof. Since $\overline{E}(1_A) = 1$, there exists an IEF $E' \subseteq E$ such that $\text{supp}(E') \subseteq A$. Denote $E'_n = E' T'^n$. By Proposition 17, the limit $M'_A := \lim_{n \rightarrow \infty} E'_n$ exists and is the unique such limit IEF corresponding to T' . Moreover $M'_A = M'_A T' \subseteq M'_A T$ and therefore, the limit $M_A = \lim_{n \rightarrow \infty} M'_A T^n$ exists as well, by Proposition 2, and is T -invariant. We will show that it satisfies (14).

The convergence of the sequence $\{E'_n\}$ to M'_A implies that for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\overline{E}'_n(f + \varepsilon 1_{\mathcal{X}}) = \overline{E}'_n(f) + \varepsilon \geq \overline{M}'_A(f)$ for every

$n \geq N$ and every $f \in \mathcal{L}_1$, which is a compact set of gambles. Further, we denote $E_n = ET^n$, and clearly $E_n \supseteq E'_n$ holds. Hence,

$$\overline{E}_{N+k}(f) + \varepsilon = \overline{E}_{N+k}(f + \varepsilon 1_{\mathcal{X}}) \geq \overline{E}'_{N+k}(f + \varepsilon 1_{\mathcal{X}}) = \overline{E}'_{N+k}(f) + \varepsilon \geq \overline{M}'_A(T^k f).$$

Letting $k \rightarrow \infty$ on both sides we obtain that

$$\liminf_{n \rightarrow \infty} \overline{E}_n(f) + \varepsilon \geq \overline{M}_A(f)$$

for every $\varepsilon > 0$, and therefore $\liminf_{n \rightarrow \infty} \overline{E}_n(f) \geq \overline{M}_A(f)$ for every $f \in \mathcal{L}_1$, and therefore also for every gamble $f \in \mathcal{L}$. \square \square

Corollary 8. *Let A, T' and M_A satisfy conditions as in Corollary 7, and let E be an IEF such that*

$$\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n 1_A) = 1.$$

Then

$$\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) \geq \overline{M}_A(f) \quad (15)$$

for every gamble f .

Proof. Let us first show that, for some constant α , $\overline{F}(1_A) \geq \alpha$ implies that

$$\liminf_{n \rightarrow \infty} \overline{F}(\overline{T}^n f) \geq \alpha \overline{M}_A(f) \quad (16)$$

for every gamble f . By Corollary 7 we have that

$$\liminf_{n \rightarrow \infty} P(\overline{T}^n f) \geq \overline{M}_A(f) \quad (17)$$

for every (imprecise) expectation functional P such that $P(1_A) = 1$. Now suppose that $\overline{F}(1_A) \geq \alpha$. Then there exists some $P \in F$ such that $P(1_A) \geq \alpha$, and by definition we have that $P(\overline{T}^n f) \leq \overline{F}(\overline{T}^n f)$ for every $n \in \mathbb{N}$. Moreover, we have that

$$P(\overline{T}^n f) = P(1_A)P(\overline{T}^n f|A) + P(1_{A^c})P(\overline{T}^n f|A^c),$$

where $P(\cdot|A)$ is the conditional expectation functional. Hence,

$$P(\overline{T}^n f) \geq \alpha P(\overline{T}^n f|A) = \alpha P(\cdot|A)(\overline{T}^n f).$$

However, by (17), $\liminf_{n \rightarrow \infty} P(\cdot|A)(\overline{T}^n f) \geq \overline{M}_A(f)$, whence

$$\liminf_{n \rightarrow \infty} \overline{F}(\overline{T}^n f) \geq \liminf_{n \rightarrow \infty} P(\overline{T}^n f) \geq \alpha \overline{M}_A(f).$$

and this proves (16). Now, for every $\varepsilon > 0$ there is some $n \in \mathbb{N}$ such that $\overline{E}(\overline{T}^n 1_A) > 1 - \varepsilon$. Hence, for every gamble f ,

$$\liminf_{k \rightarrow \infty} \overline{E}(\overline{T}^n \overline{T}^k f) \geq (1 - \varepsilon) \overline{M}_A(\overline{T}^n f) = (1 - \varepsilon) \overline{M}_A(f).$$

Since the above holds for every $\varepsilon > 0$, inequality (15) follows. \square \square

Corollary 9. *Let B be a minimal permanent class. Then a T -invariant IEF M_B exists such that for any IEF E with*

$$\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n 1_B) = 1. \quad (18)$$

the inequality

$$\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) \geq \overline{M}_B(f) \quad (19)$$

holds for every gamble f .

Proof. By Proposition 11, B is absorbing and regular with respect to $(T_B^r)^k \subseteq T^{rk}$. Therefore, by Corollary 8, IEF M_B exists such that

$$\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^{nkr} f) \geq \overline{M}_B(f)$$

for every IEF E satisfying (18), and therefore also for ET^m , where m is arbitrary integer. Thus we have

$$\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^{m+nkr} f) = \liminf_{n \rightarrow \infty} \overline{ET}^m(\overline{T}^{nkr} f) \geq \overline{M}_B(f).$$

It remains to prove that M_B is T -invariant. By Corollary 7, it is T^{rk} -invariant. Therefore $M_B T = M_B T^{nkr+1}$. Replacing E with M_B in (19), and using the above identity imply that $M_B(Tf) \geq M_B(f)$ for every gamble, which implies the inclusions $M_B \subseteq M_B T \subseteq \dots \subseteq M_B T^{kr} = M_B$. But then all the inclusions must in fact be equalities, particularly, $M_B = M_B T$ must hold. $\square \quad \square$

Theorem 3. *Let E be an IEF on \mathcal{X} such that the limit*

$$\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n 1_B) =: \overline{E}_\infty(1_B)$$

exists for every minimal permanent class B , and $\overline{E}_\infty(1_B) \in \{0, 1\}$. Then the limit

$$\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) =: \overline{E}_\infty(f)$$

exists for every gamble f . Moreover, E_∞ is the unique T -invariant IEF such that $\overline{E}_\infty(1_B) = 1$ for every minimal permanent class $B \subseteq S_E$.

Proof. Let B be a minimal permanent class such that

$$\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n 1_B) = 1. \quad (20)$$

Then by Corollary 9 we have that $\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) \geq \overline{M}_B(f)$. Let us show that this holds for every minimal permanent class $B \subseteq S_E$. Suppose contrary that there is some $B \subseteq S_E$ such that $\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n 1_B) \neq 1$. Then, by the assumption, $\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n 1_B) = 0$. But, by definition of S_E , there must be some $m \geq 0$ such that $\overline{E}(\overline{T}^m 1_B) > 0$, which by Proposition 15 implies that $\overline{E}(\overline{T}^{m+n} 1_B) \geq \overline{E}(\overline{T}^m 1_B)$, and therefore the limit as $n \rightarrow \infty$ cannot be 0. This contradiction confirms that (20) holds for every minimal permanent class $B \subseteq S_E$.

By the above, for every minimal permanent class $B \subseteq S_E$ we have that $\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) \geq \overline{M}_B(f)$. Let us define

$$\overline{M}_E(f) = \max_{B \subseteq S_E} \overline{M}_B(f),$$

which is an IEF with $\overline{M}_E(1_B) = 1$ for every $B \subseteq S_E$. Moreover, since

$$\overline{M}_E(\overline{T}f) = \max_{B \subseteq S_E} \overline{M}_B(\overline{T}f) = \max_{B \subseteq S_E} \overline{M}_B(f) = \overline{M}_E(f).$$

it is a T -invariant IEF, and by Theorem 2, it must therefore be the unique T -invariant IEF with these properties on S_E . Hence,

$$M_E = \lim_{n \rightarrow \infty} V_E T^n,$$

where V_E is the vacuous IEF on S_E . For the vacuous IEF we have that $E \subseteq V_E$, and therefore, for every gamble f we have that

$$\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) \geq \overline{M}_E(f) = \lim_{n \rightarrow \infty} \overline{V}_E(\overline{T}^n f) \geq \limsup_{n \rightarrow \infty} \overline{E}(\overline{T}^n f).$$

Hence, $\liminf_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) = \limsup_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) = \lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n f) = \overline{M}_E(f)$. \square

Let us conclude this section with an example.

Example 5. Let \mathcal{X} be a set of states partitioned into $\mathcal{X}_1, \mathcal{X}_2$ and \mathcal{X}_3 . Suppose that an ITO can be decomposed in the following matrix form:

$$T = \begin{pmatrix} T_{11} & 0 & 0 \\ T_{21} & T_{22} & 0 \\ T_{31} & T_{32} & T_{33} \end{pmatrix}$$

where $\overline{T}_{ii}1_{\mathcal{X}_i}(x) = 1$ for every $i = 1, 2, 3$ and every $x \in \mathcal{X}_i$. Moreover, assume that T induces a regular weak accessibility relation on every \mathcal{X}_i . Then \mathcal{X}_i are all minimal permanent classes. If $\overline{T}_{ij} = 0$ whenever $i \neq j$, then we can have that $\overline{E}(1_{\mathcal{X}_i})$ equals 0 or 1 independently of each other, with the only condition that $\overline{E}(1_{\mathcal{X}}) = 1$. Therefore we have seven different extremal T -invariant IEFs.

However, in the case where $\overline{T}_{ij} > 0$ for some $i \neq j$, $\overline{E}(1_{\mathcal{X}_i})$ clearly implies that $\lim_{n \rightarrow \infty} \overline{E}(\overline{T}^n 1_{\mathcal{X}_j}) = 1$ as well, which restricts the number of possible distinct extremal IEFs.

In addition to the extremal T -invariant IEFs, we can obtain the non-extremal ones by forming convex combinations. It is one of the challenges left to explore, whether every non-extremal T -invariant IEF is equal to such a convex combination. If the answer is positive, this would allow us to characterise all T -invariant IEFs.

5 Conclusions

The first main contribution of this paper is the introduction of the so-called strong accessibility relation which together with the weak relation, defined earlier by de Cooman et al. [2009], allows a detailed analysis of the behaviour of imprecise Markov chains. We have identified the minimal sets of states, named the minimal permanent classes, that correspond to essential communication classes in the classical theory. They are, similarly, the minimal sets of states where the entire probability mass can be concentrated and possibly remains there forever.

Once the minimal irreducible building blocks had been identified, we could describe and analyse the invariant distributions whose probability mass is entirely concentrated within the minimal permanent classes, and this is also the point of the crucial difference between precise and imprecise theory. While in the precise theory, the entire probability mass can be in one class at the time only, it can 'possibly' be in several classes when the 'imprecise interpretation' is adopted. Therefore the behaviour of an imprecise Markov chain cannot be completely reduced to the behaviour within minimal permanent classes. Instead, we study the so-called extremal imprecise expectation functionals, which allow the probability mass to be concentrated entirely in one or more permanent classes at the time.

The extremal imprecise expectation functionals seem to be the minimal objects in the class of imprecise expectation functionals that cannot be further reduced to simpler constituent parts. Our main convergence result in Theorem 2 then shows that when an invariant imprecise distribution is extremal then it is uniquely determined with the corresponding set of possible minimal permanent classes. Moreover, as shown in Theorem 3, the unique convergence to extremal invariant distributions is ensured already with the convergence of the restrictions to the characteristic functions of minimal permanent classes to either 0 or 1.

Convex combinations of extremal invariant imprecise expectation functionals are the limits of the same combinations of the initial imprecise expectation functionals. Two important challenges that remain subject to further work arise. First is to explore whether every invariant imprecise expectation functional is only a convex combination of a family of extremal ones, and the related question is whether unique convergence when restricted to minimal permanent classes is enough for unique convergence of imprecise Markov chain. Another challenge is the analysis of periodic behaviour in imprecise Markov chains.

References

- R.J. Crossman and D. Škulj. Imprecise Markov chains with absorption. *International Journal of Approximate Reasoning*, 51:1085–1099, 2010. ISSN 0888-613X. doi: <http://dx.doi.org/10.1016/j.ijar.2010.08.008>.
- G. de Cooman, F. Hermans, and E. Quaeghebeur. Imprecise Markov chains and their limit behavior. *Probability in the Engineering and Information Sciences*, 23(4):597–635, 2009. ISSN 0269-9648. doi: {10.1017/S0269964809990039}.
- D.J. Hartfiel. *Markov Set-Chains*. Springer-Verlag, Berlin, 1998.
- D.J. Hartfiel and E. Seneta. On the theory of Markov set-chains. *Advances in Applied Probability*, 26(4):947–964, 1994.
- E. Seneta. *Non-Negative Matrices and Markov Chains*. Springer, 2006.
- D. Škulj. Discrete time Markov chains with interval probabilities. *International Journal of Approximate Reasoning*, 50(8):1314–1329, 2009. ISSN 0888-613X. doi: <http://dx.doi.org/10.1016/j.ijar.2009.06.007>.

- D. Škulj and R. Hable. Coefficients of ergodicity for Markov chains with uncertain parameters. *Metrika*, pages 1–27, 2012. ISSN 0026-1335. URL <http://dx.doi.org/10.1007/s00184-011-0378-0>. 10.1007/s00184-011-0378-0.
- P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, New York, 1991.
- K. Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung. I: Intervallwahrscheinlichkeit als umfassendes Konzept*. Physica-Verlag, Heidelberg, 2001.